



Comunidade de Suporte da Cisco -
Webcast ao vivo:

Border Gateway Protocol (BGP) – Fundamentos, Configuração e Troubleshooting

Marcio Ferreira

Terça-feira 18 de junho de 2013

Webcast com Especialistas em Tecnologia da Comunidade Cisco

Especialista de hoje:

Marcio Ferreira, HTE do Focused Technical Services da Cisco Brasil.



Marcio Ferreira

Webcast com Especialistas em Tecnologia da Comunidade Cisco

Especialista ajudante de hoje:

Rafael Lima, Engenheiro HTTS do TAC da Cisco Brasil.



Rafael Lima

Obrigada por estar com a gente hoje!

Durante a apresentação, serão feitas
algumas perguntas para o público.

Dê suas respostas, participe!



Primeira Pergunta

Qual é sua experiência com BGP?

- a) Básica. Já tive alguns contatos, porém não entendo muito sobre a solução em si.
- b) Eu tenho conhecimento avançado, porém utilizo o time do TAC em muitos casos.
- c) Estou em processo de aprendizado.
- d) Não tenho ideia sobre esta solução.



Webcast com Especialistas em Tecnologia da Comunidade Cisco:

Border Gateway Protocol (BGP) – Fundamentos, Configuração e Troubleshooting

Marcio Ferreira

Data: 18/06/2013

Agenda

- Objetivos
- Conceitos Básicos
- Como funciona?
- Vizinhanças
- Geração de Prefixos
- Seleção do Melhor Caminho

Conceitos Básicos

Sistema Autônomo (AS)

- Uma coleção de prefixos sobre uma mesma política de roteamento e sob controle administrativo de uma mesma entidade

Pode ser uma rede extensa ou pequena

Todo prefixo anunciado na Internet deve pertencer a um AS

- Conectividade interna contígua
- Numeração válida entre 1 e 65,535 (2-byte ASN)

Range privado: 64512 – 65535

- Range estendido (4-byte ASN)

O que é um IGP?

- Interior Gateway Protocol
- Protocolo utilizado dentro de um Sistema Autônomo (AS)
- A principal função de um IGP é convergência rápida da rede e que todos os roteadores possuam visão completa da topologia de rede
- Carrega informação detalhada sobre prefixos internos
- Exemplos — OSPF, ISIS, EIGRP, RIP, etc..

O que é um EGP?

- Exterior Gateway Protocol
- Utilizado para troca de informação de rotas entre Sistema Autônomos
- Disassociado do IGP
- EGP utilizado atualmente na Internet é o BGP versão 4

Por que é necessário um EGP?

- Escalabilidade
 - Hierarquia
 - Isolar flooding periódico de IGP's
 - Isolar instabilidades de rede
- Suporte a políticas de Rede Complexas
 - Controlar acessibilidade a prefixos
 - Unir organizações separadas
 - Conectar múltiplos IGP's

Roteamento Interior vs. Exterior

- Interior

Descobrimento automático de vizinhos

Roteadores internos possuem informação completa da tabela de rotas

- Exterior

Vizinhos configurados estaticamente

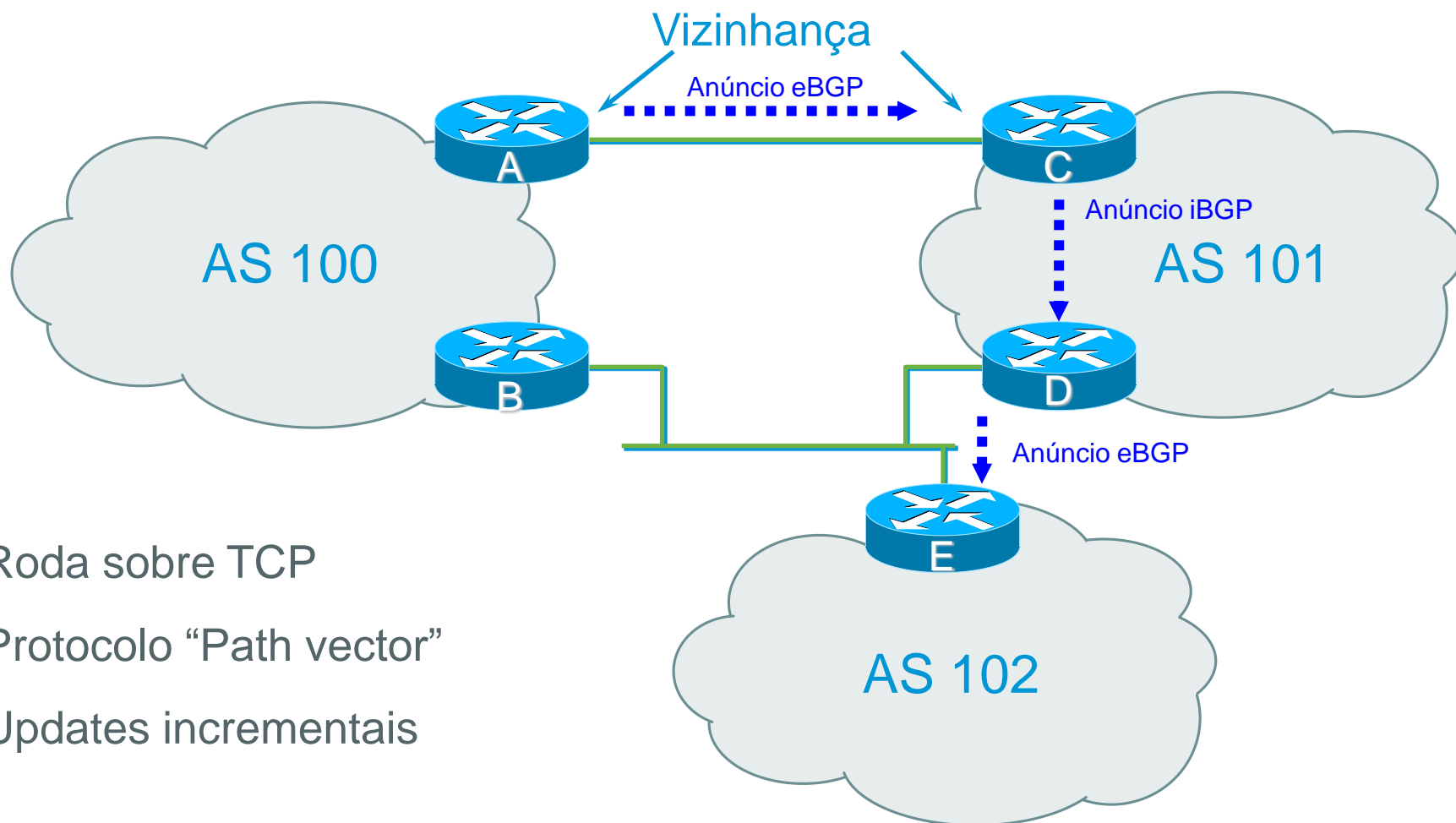
Conexão com redes externas

Demarcação de limites administrativos

Como funciona?



BGP – Como funciona



- Roda sobre TCP
- Protocolo “Path vector”
- Updates incrementais

Modo de Operação Geral

- Roteadores BGP aprendem múltiplos caminhos via vizinhos BGP internos e externos.
- Escolhe SOMENTE o melhor caminho e instala na tabela de roteamento IP.
- BGP é o protocolo com maior suporte a políticas de filtragem de rotas e manipulação de atributos para influenciar no processo de escolha do melhor caminho.

Descrição - rfc1771

- Roteador BGP anuncia apenas as rotas que este utiliza

Paradigma de roteamento “hop-by-hop”

- Protocolo de Transporte Confiável

Não há necessidade de implementar fragmentação, reTX, ACKs e sequenciação

assume um fechamento “graceful”: toda informação será entregue

Resumo de Operação

- Sessão utiliza conexão TCP (porta 179)
- Troca da tabela completa após inicialização da vizinhança
- Updates incrementais
- Somente a melhor rota é utilizada
- Somente a rota utilizada é anunciada

Segunda Pergunta

Quantos caminhos o BGP pode considerar como melhor caminho para um prefixo?

- a) 1
- b) 2
- c) 3
- d) **Depende da configuração de Multipath BGP**

Vizinhanças



Vizinhanças Externas (eBGP)

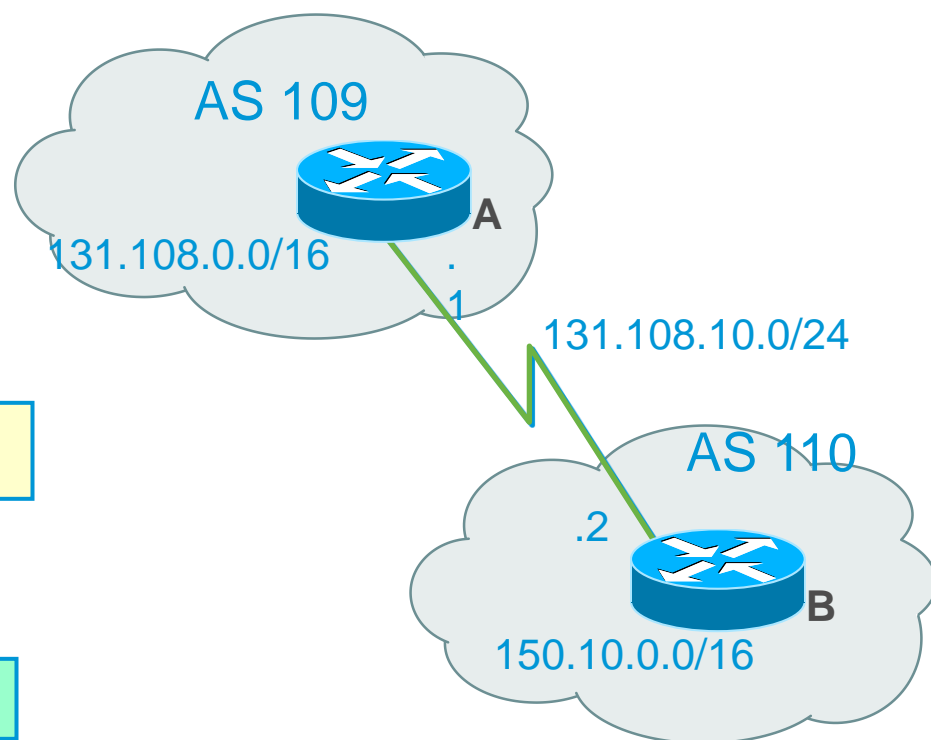
- Vizinhos BGP em AS diferentes
- Devem estar diretamente conectados
- Configuração:

Router B

```
router bgp 110  
network 150.10.0.0  
neighbor 131.108.10.1 remote-as 109
```

Router A

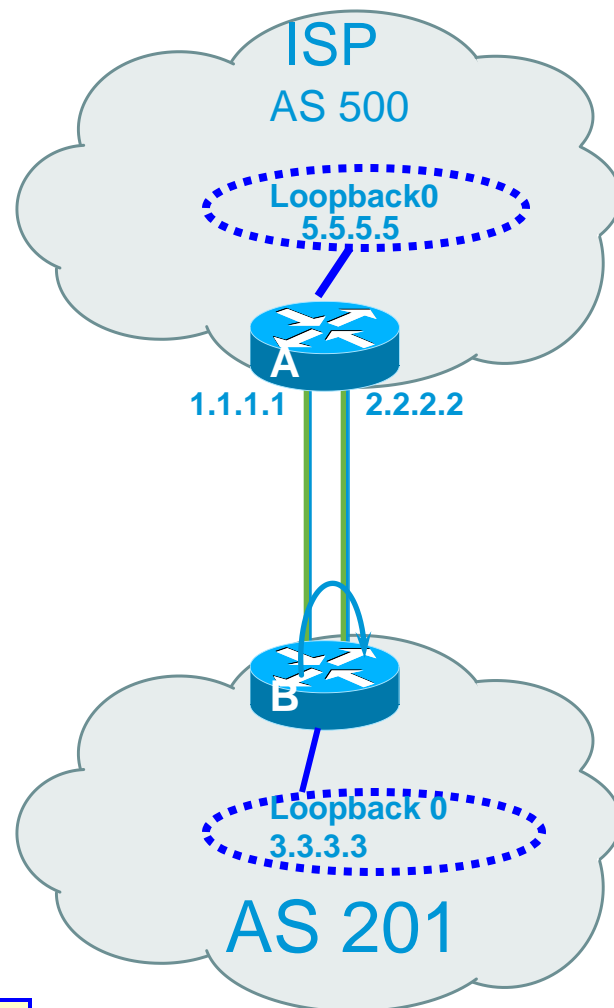
```
router bgp 109  
network 131.108.0.0  
neighbor 131.108.10.2 remote-as 110
```



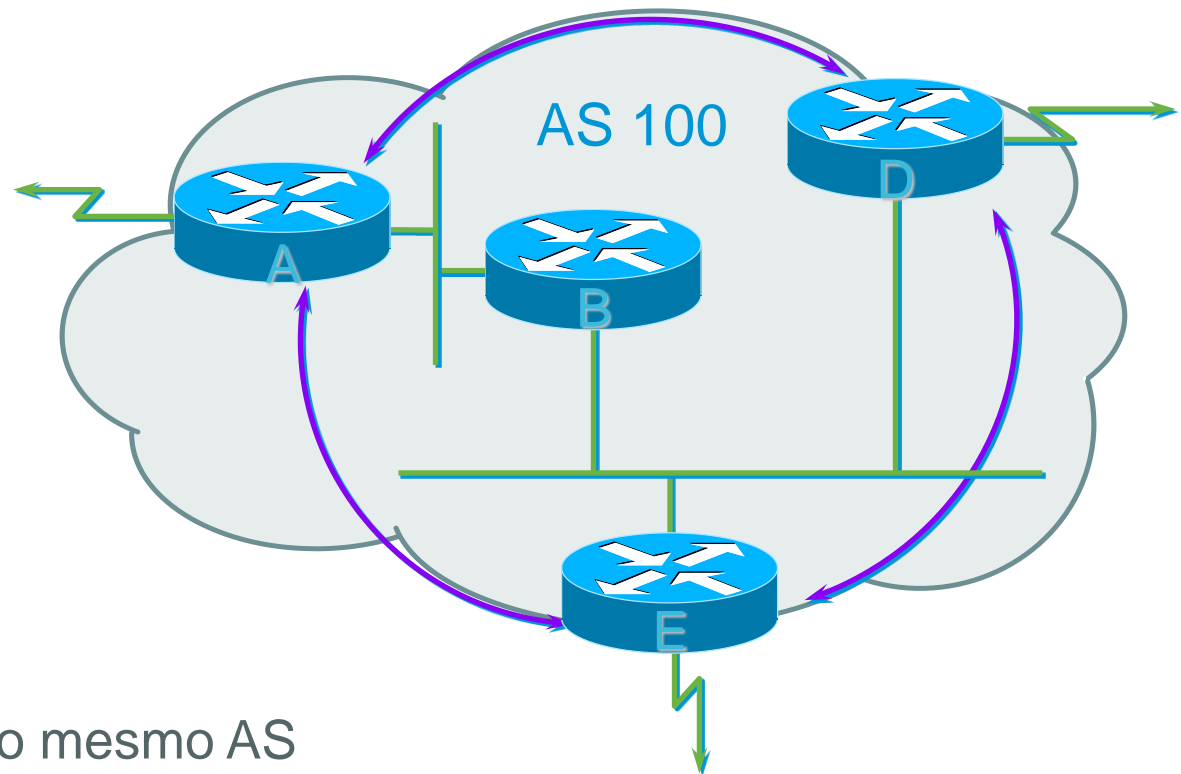
Links Paralelos

- Utilizar `<ebgp-multihop>`
- Configurar interfaces loopbacks nos dois roteadores
- Configurar um IGP (geralmente rota estática) entre interfaces loopbacks
- Configuração roteador B:
router bgp 201
neighbor 5.5.5.5 remote-as 500
neighbor 5.5.5.5 update-source loopback0
neighbor 5.5.5.5 ebgp-multihop
!
ip route 5.5.5.5 255.255.255.255 1.1.1.1
ip route 5.5.5.5 255.255.255.255 2.2.2.2

Valor Default = 255 caso
não especificado



Vizinhanças Internas (iBGP)



- Vizinhos BGP dentro do mesmo AS
- Não precisam estar diretamente conectados
- Vizinhos iBGP devem falar todos BGP entre si (full mesh)

Vizinhanças Internas (iBGP)

- Configuração:

Router B

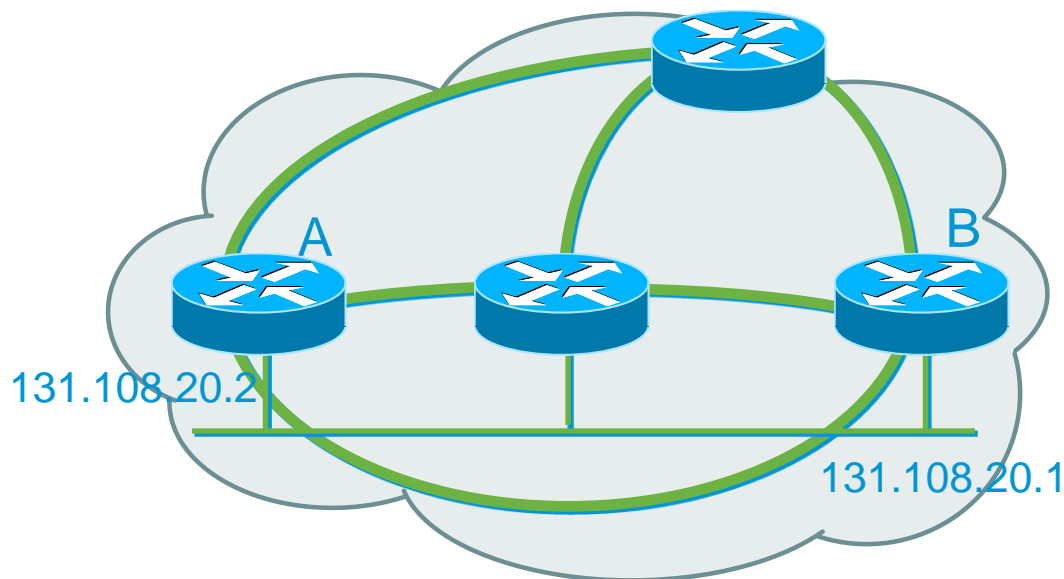
```
router bgp 109
```

```
neighbor 131.108.20.2 remote-as 109
```

Router A

```
router bgp 109
```

```
neighbor 131.108.20.1 remote-as 109
```



Vizinhanças Internas (iBGP)

- Utilizando `<update-source>`

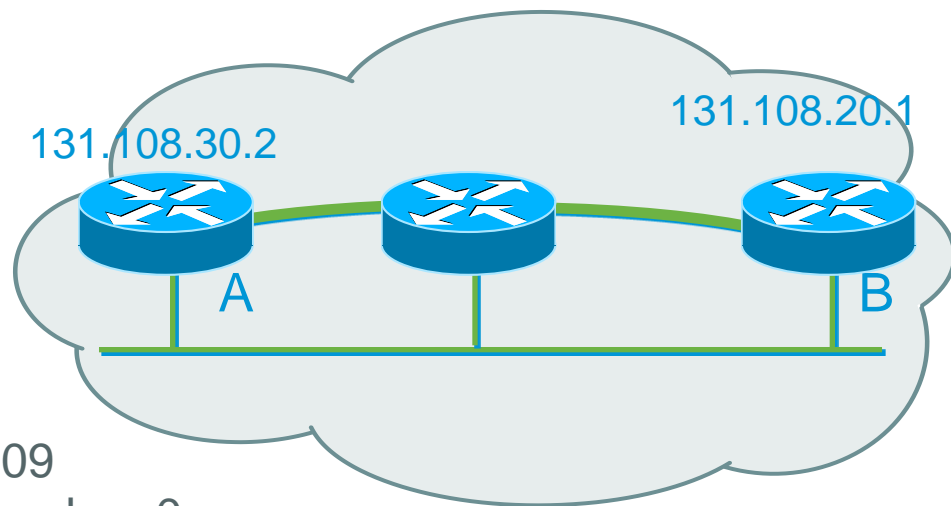
- Configuração:

Router B

```
router bgp 109  
neighbor 131.109.30.2 remote-as 109  
neighbor 131.109.30.2 update-source loop0
```

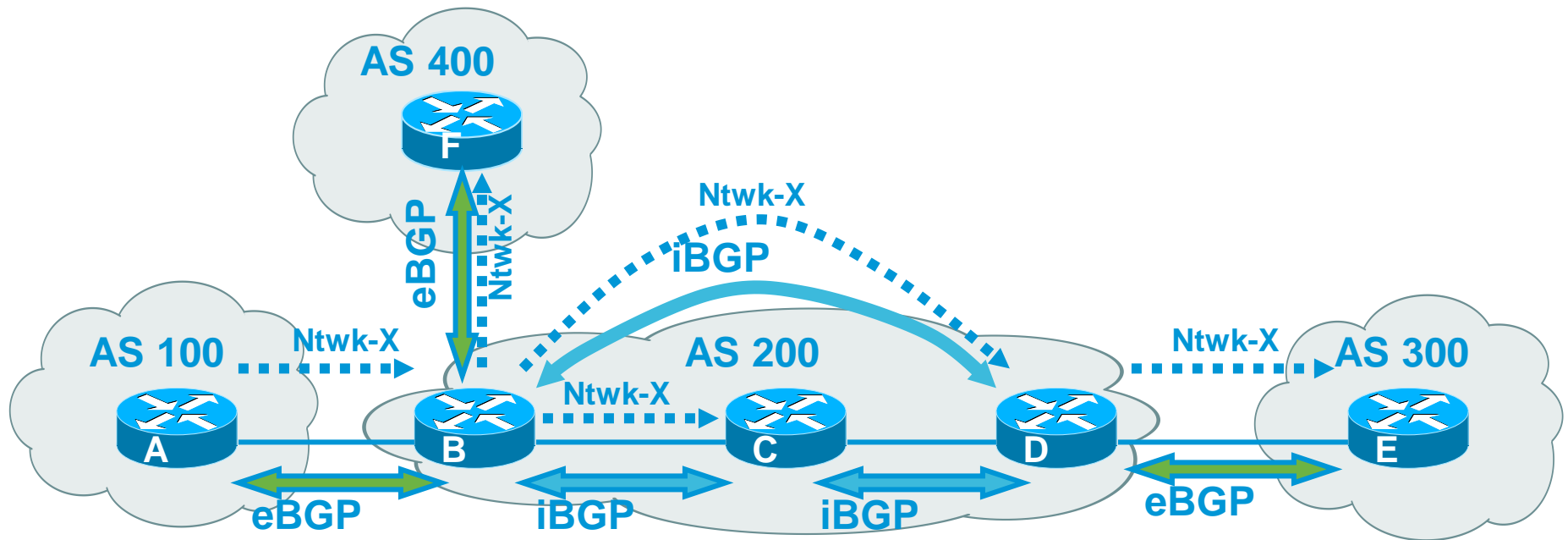
Router A

```
router bgp 109  
neighbor 131.108.20.1 remote-as 109  
neighbor 131.108.20.1 update-source loop0
```



Transferência de Informação (eBGP)

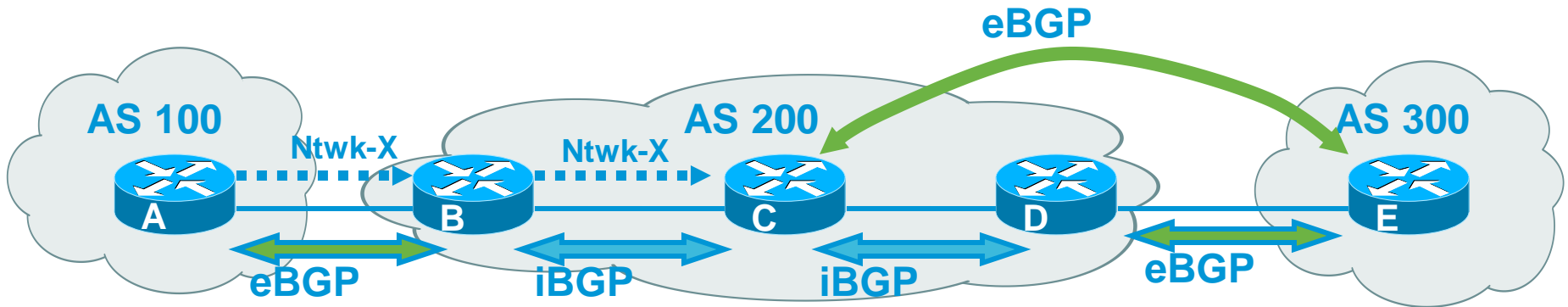
- Recebido de eBGP -> anuncia para todos vizinhos iBGP e eBGP



- Propaga **SOMENTE** a melhor rota

Transferência de Informação (iBGP)

- Recebido de iBGP -> anuncia para vizinhos eBGP
Requerido full mesh iBGP!!



- Propaga **SOMENTE** a melhor rota

NEXT_HOP

- O next hop para alcançar uma rede

eBGP

endereço IP do vizinho

iBGP

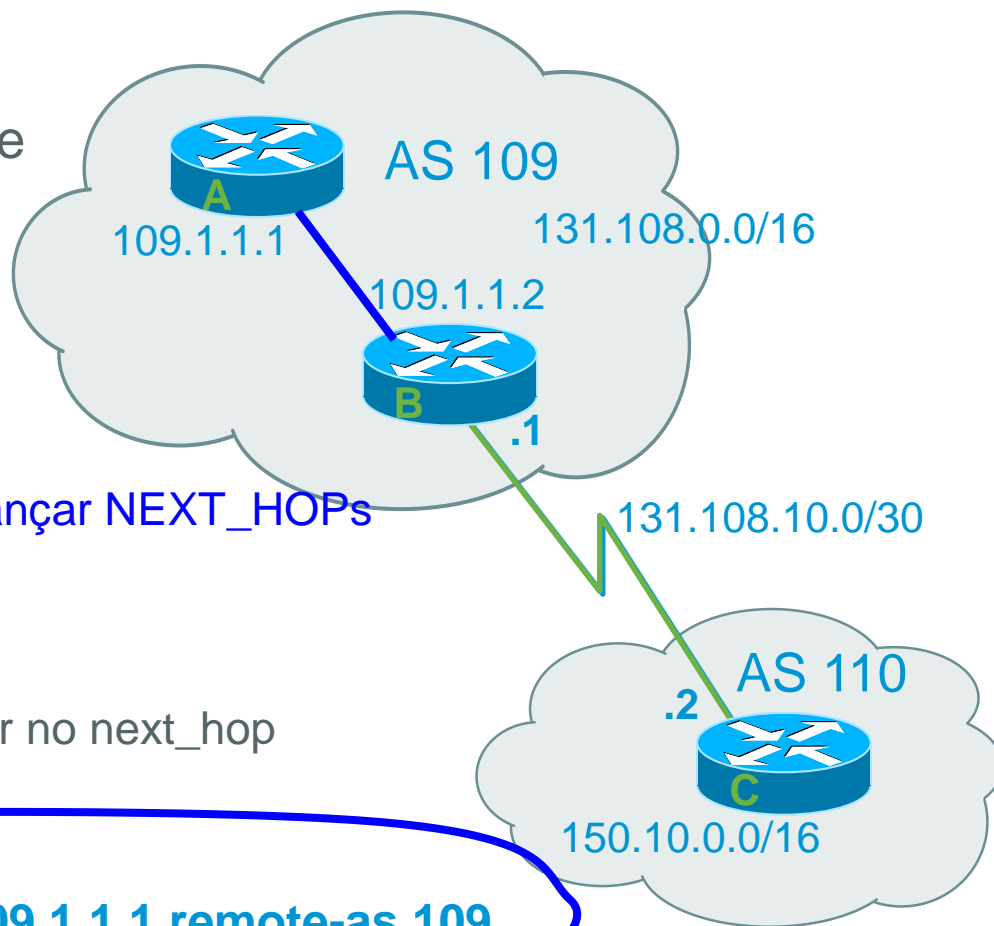
NEXT_HOP anunciado pelo eBGP

IGP carrega informação de como alcançar NEXT_HOPs

“route lookup” recursivo

Desassocia BGP da topologia física

Permite ao IGP controlar como chegar no next_hop



Ou...

```
Router-B(config-router)# neighbor 109.1.1.1 remote-as 109
```

```
Router-B(config-router)# neighbor 109.1.1.1 next-hop-self
```

Geração de Prefixos



Gerando Anúncios BGP

1. Redistribuição dinâmica do IGP

- Nem sempre a melhor opção
 - Pode anunciar redes privadas (redes que não deveriam ser conhecidas pela Internet)
- Requer configurações cuidadosas de filtragem de rotas
- Caso um prefixo não esteja na tabela do IGP este não pode ser anunciado
- Mais utilizado em cenários PE-CE

Gerando Anúncios BGP

2. Redistribuição de rotas estáticas apontando para Null0

- Mesmo que os prefixos estejam fora do IGP as redes são anunciadas
- Configurados nos roteadores de peering Internet

- Exemplo:

```
router bgp 109
 redistribute static
 !
 ip route 198.10.0.0 255.255.0.0 null0
```

Terceira Opção...comando <network>

- Redes originadas pelo roteador local
- Rota deve existir no IGP
 - Entrada dinâmica ou estática/conectada na tabela de roteamento

Exemplo:

```
!  
router bgp 109  
network 200.10.10.0  
network 198.10.0.0 mask 255.255.0.0  
!
```


O comando <network> e a opção “mask”

- Quando não é necessária a opção “mask”?

Somente quando o prefixo especificado é classfull e existe em sua tabela de roteamento IGP.

network 10.0.0.0 → isto **irá** funcionar desde que você tenha esta rota específica em sua tabela de roteamento IGP.

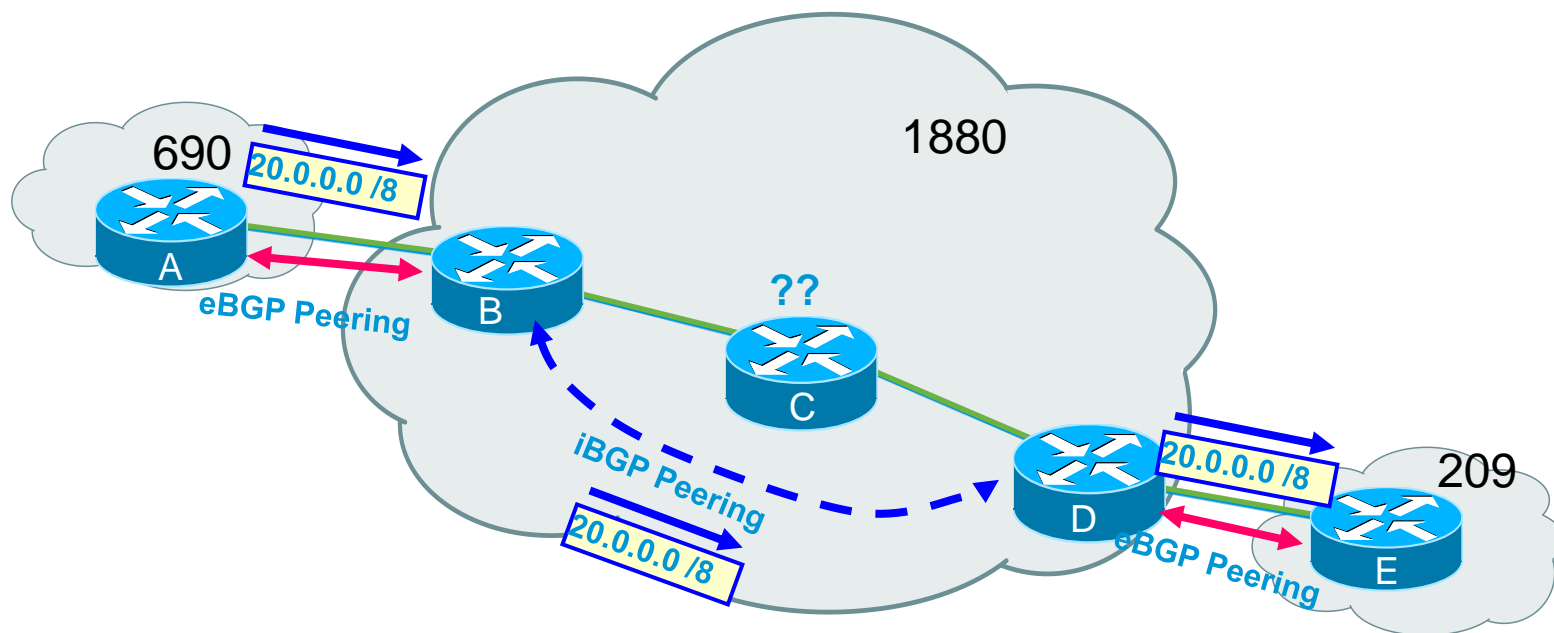
network 10.10.0.0 → isto **não** irá funcionar... mesmo que você tenha esta rota específica em sua tabela de roteamento IGP.

network 10.10.0.0 mask 255.255.0.0 → isto **irá** funcionar desde que você tenha esta rota específica em sua tabela de roteamento IGP.

Sincronização

- Em uma rede transito, uma rota aprendida de um vizinho externo não deve ser anunciada para outros vizinhos eBGP até que todos os roteadores no AS local tenham aprendido esta rota.

Porque a necessidade de Sincronização



O que aconteceria com Sincronização configurada como **OFF**

1. Router-A anuncia rede 20.0.0.0/8 via eBGP para Router-B.
2. Router-B anuncia esta mesma rota **via iBGP** para Router-D.
3. Router-C não fala BGP então não aprende esta rota.
4. **Router-D anuncia a rota via eBGP para o Router-E.** → **Este passo não ocorreria com Sincronização.**
5. Se pacotes do AS-209 chegam no AS-1880 com um IP de destino 20.x.x.x estes serão dropados pelo Router-C. **Rede de transito quebrada!**

no synchronization

Sincronização é desligada por default no IOS atualmente

Por que desligar?

- rede não faz transito para outras redes externas
- todos os roteadores no caminho rodam BGP

Vantagens de desligar?

- menos rotas no IGP
- convergência mais rápida do BGP(não precisa perder tempo checando tabela do IGP)

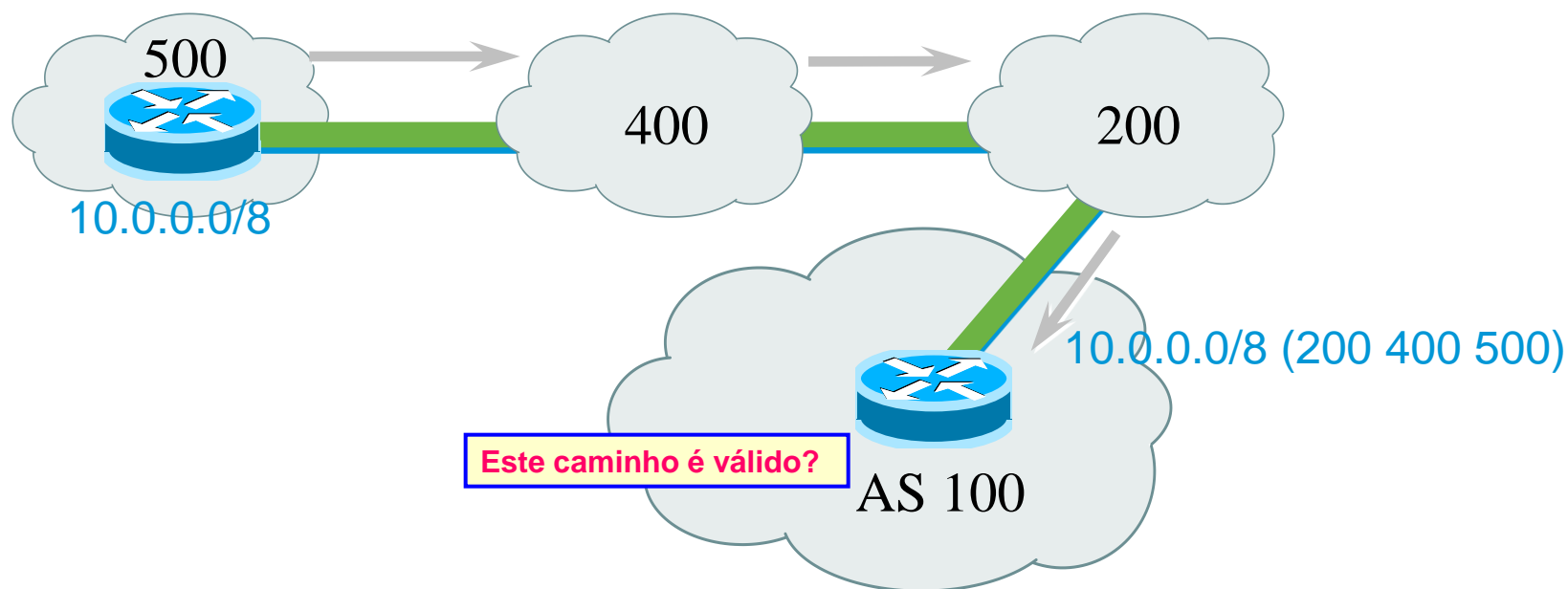
Seleção do Melhor Caminho



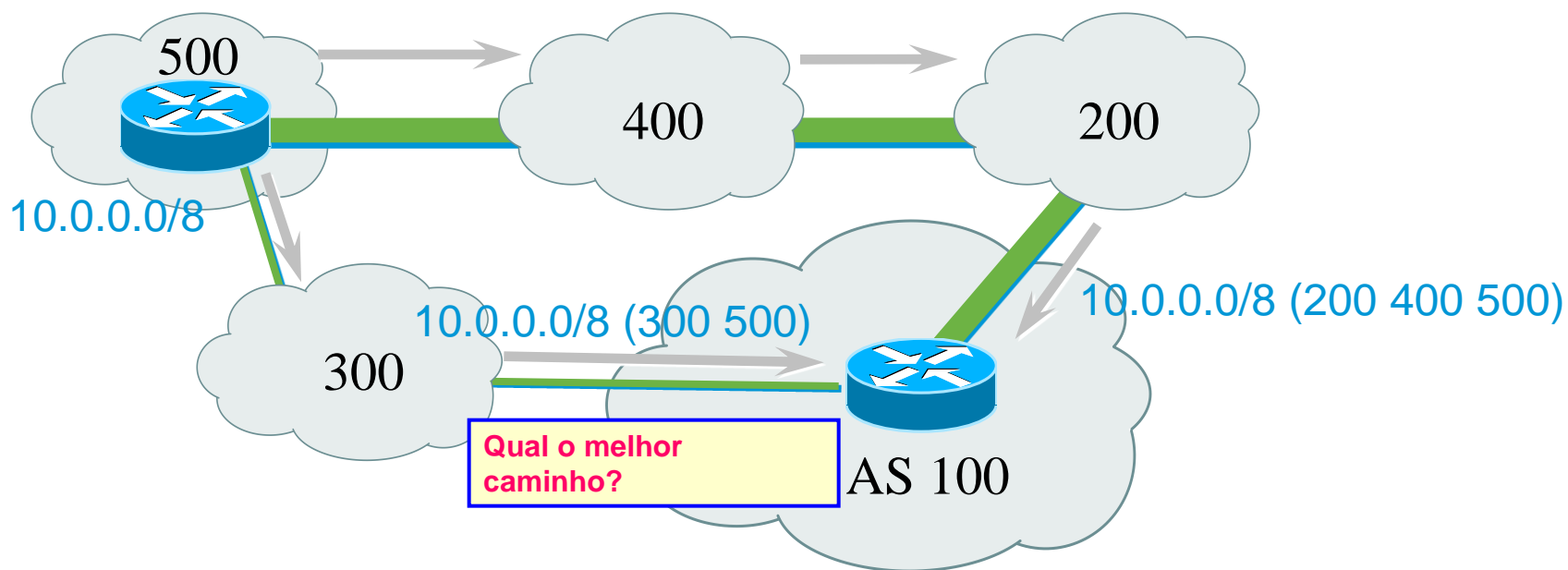
Atributos de Rotas BGP

- Todo prefixo recebido via BGP possui Atributos, que são utilizados para escolha do melhor caminho. Estes atributos também podem ser configurados localmente e aplicados ao recebimento do prefixo.
- Alguns Atributos devem ser sempre enviados quando gerado um update BGP -> **AS_PATH, ORIGIN, NEXT_HOP**
- Alguns Atributos não devem ser enviados (são significantes localmente apenas).
- Alguns atributos recebidos podem ou não ser encaminhados aos vizinhos BGP.

BGP – Escolhendo o melhor caminho

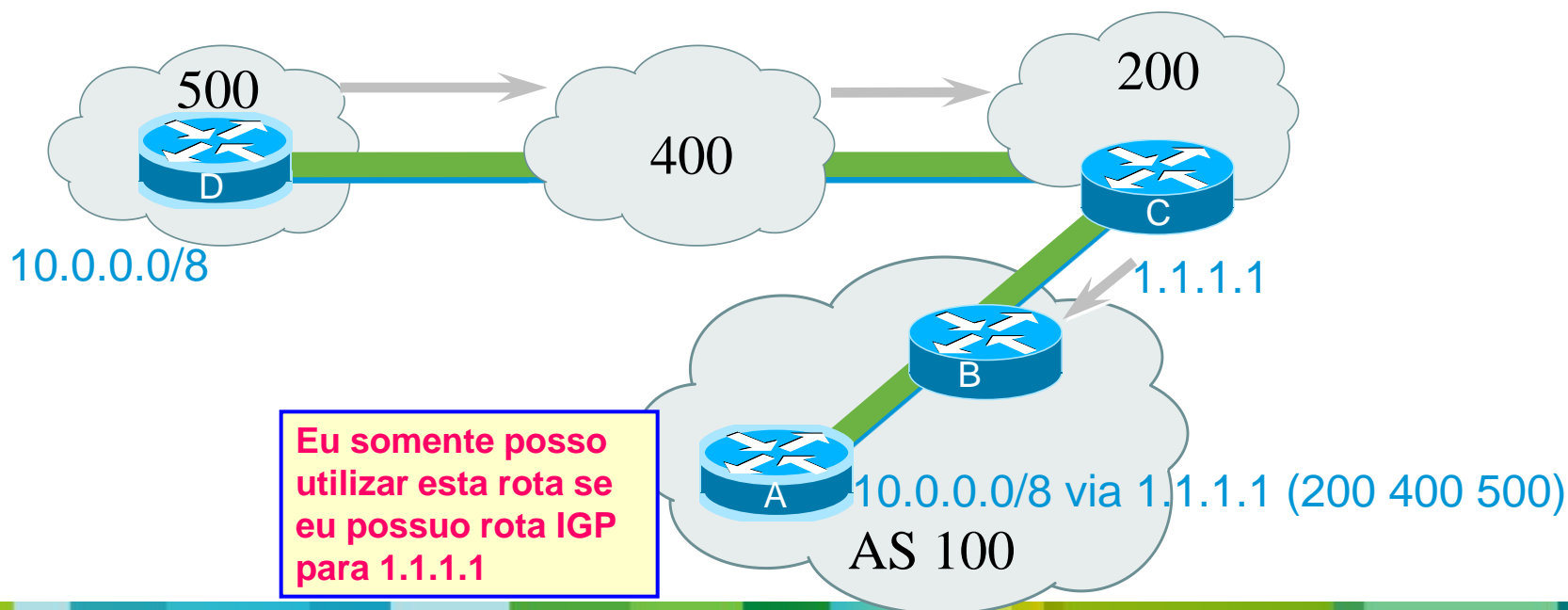


BGP – Escolhendo o melhor caminho



BGP Escolha de Melhor Caminho

- 1 Somente considerar rotas com NEXT_HOPs válidos → Well-Known Mandatory
- 2 Não considerar caminho iBGP caso não esteja sincronizado



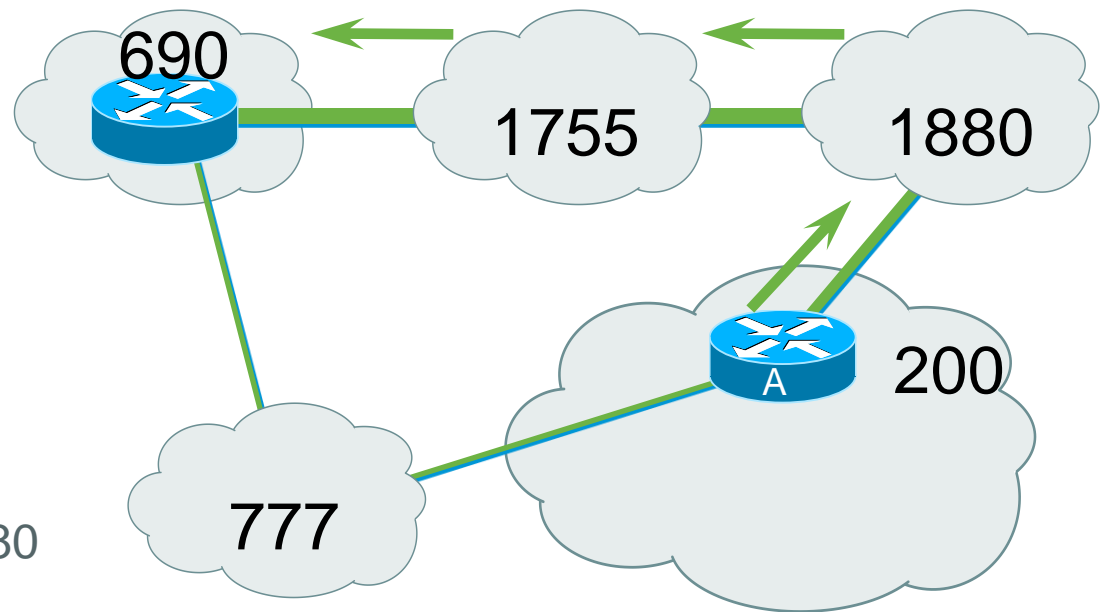
BGP Escolha de Melhor Caminho

- 1 Somentente considerar rotas com válidos
NEXT_HOPs
- 2 Não considerar caminho iBGP caso não esteja
sincronizado
- 3 **Maior WEIGHT**

WEIGHT

- Proprietário Cisco
- Local ao roteador: Não propagado
- valor: 0 – 65535
- Weight maior é preferido sobre weight menor
- Default:
 - originado localmente = 32768
 - outro = 0
- Pode ser configurado via múltiplas maneiras como route-maps, filter-lists, comando neighbor, etc


WEIGHT (Cont.)



- Configuração (rtr A):

```
!  
router bgp 200  
neighbor x.x.x.x remote-as 1880  
neighbor x.x.x.x weight 200  
!
```

BGP Escolha de Melhor Caminho

- 1 Somente considerar rotas com válidos NEXT_HOPs
- 2 Não considerar caminho iBGP caso não esteja sincronizado
- 3 Maior WEIGHT
- 4 **Maior LOCAL_PREF** 

Well-Known
Discretionary

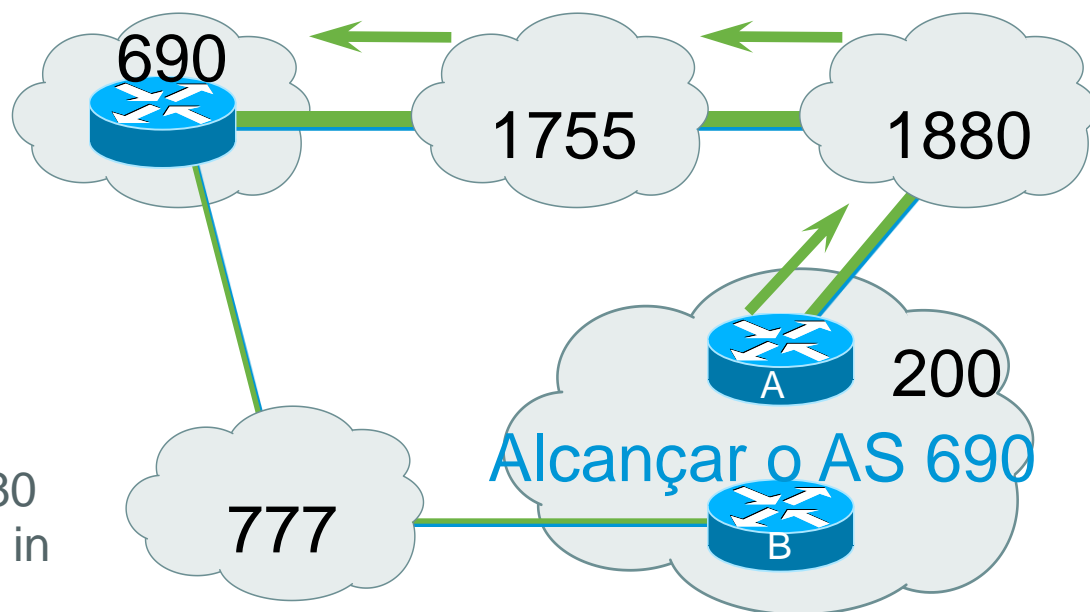
LOCAL_PREF

- Indica o caminho preferido de saída do AS local (passado para outros vizinhos BGP no mesmo AS)
- Válido globalmente no AS local
- Caminhos com maior LOCAL-PREF são mais desejáveis (default = 100)

`bgp default local-preference valor`

Não é propagado para fora do AS

LOCAL_PREF (Cont.)



- Configuração (rtr A):
router bgp 200
neighbor x.x.x.x remote-as 1880
neighbor x.x.x.x route-map foo in
!
route-map foo permit 10
match as-path 2
set local-preference 120
!
ip as-path access-list 2 permit ^1880_

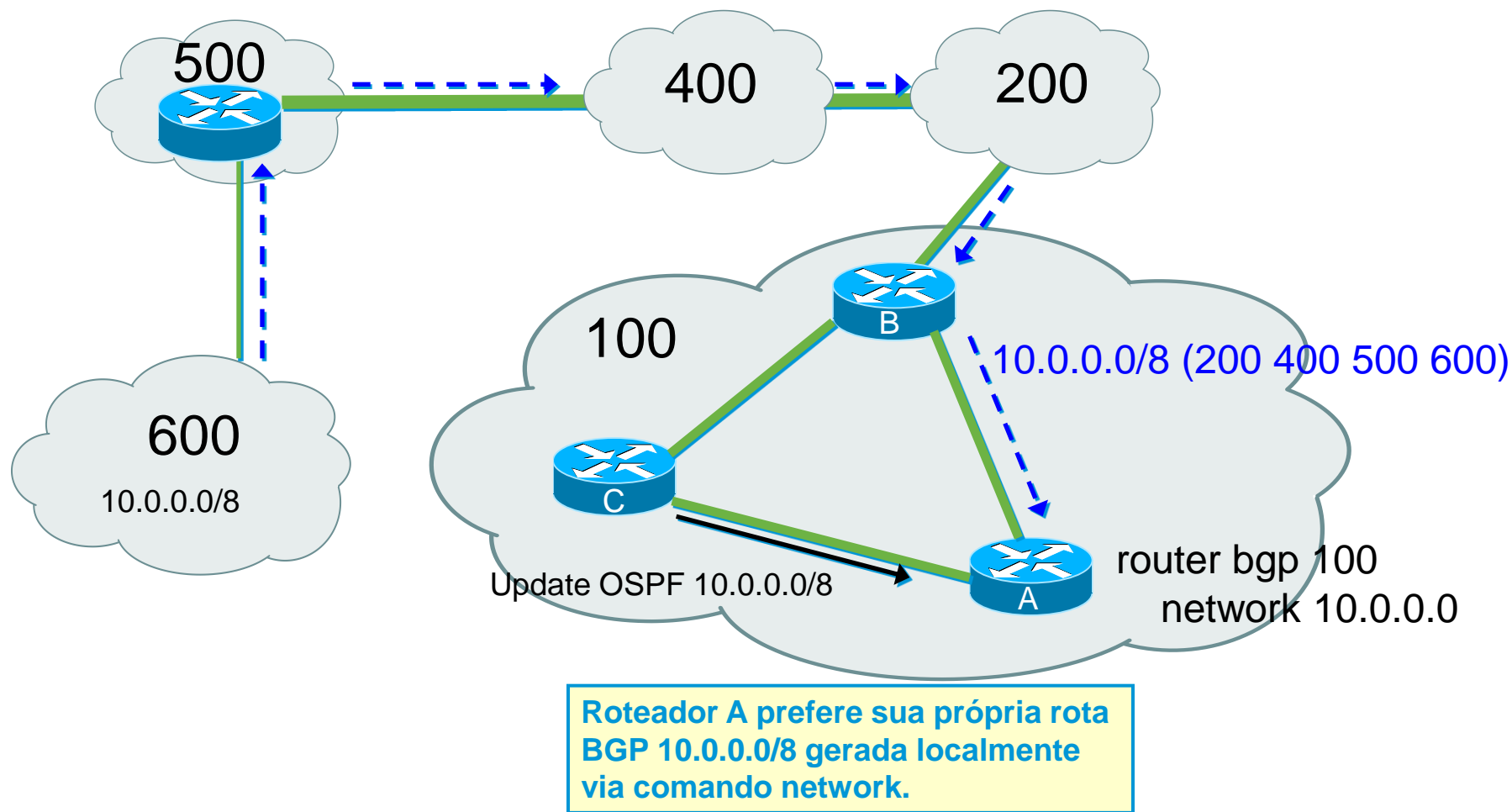
LOCAL_PREF (Cont.)

```
R3#sho ip bgp 22.0.0.0/8
BGP routing table entry for 22.0.0.0/8, version 8
Paths: (2 available, best #1, table Default-IP-Routing-Table)
Flag: 0x820
  Advertised to update-groups:
    1      2
  7
    7.7.7.7 from 7.7.7.7 (7.7.7.7)
      Origin IGP, metric 0, localpref 200, valid, external, best
  2
    2.2.2.2 from 2.2.2.2 (22.22.22.2)
      Origin IGP, metric 0, localpref 100, valid, external
R3#
```


BGP Escolha de Melhor Caminho

- 1 Somente considerar rotas com válidos NEXT_HOPs
- 2 Não considerar caminho iBGP caso não esteja sincronizado
- 3 Maior WEIGHT
- 4 Maior LOCAL_PREF
- 5 Preferir rota originada localmente

Rota originada localmente



Rota originada localmente

- R3#show run
- <output truncated>
- !
- router bgp 3
- synchronization
- bgp log-neighbor-changes
- network 22.0.0.0
- neighbor 2.2.2.2 remote-as 2
- neighbor 6.6.6.6 remote-as 3
- neighbor 7.7.7.7 remote-as 7

```
R3#sho ip bgp
BGP table version is 11, local router ID is 2.2.2.3
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
               r RIB-failure, S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network        Next Hop           Metric LocPrf Weight Path
*> 22.0.0.0       0.0.0.0             0         32768    i
*                 7.7.7.7             0         200      0 7 i
*                 2.2.2.2             0          0      0 2 i
*> 33.0.0.0       2.2.2.2             0          0      0 2 ?
R3#_
```

BGP Escolha de Melhor Caminho

- 1 Somente considerar rotas com válidos NEXT_HOPs
- 2 Não considerar caminho iBGP caso não esteja sincronizado
- 3 Maior WEIGHT
- 4 Maior LOCAL_PREF
- 5 Preferir rota originada localmente
- 6 Menor AS_PATH

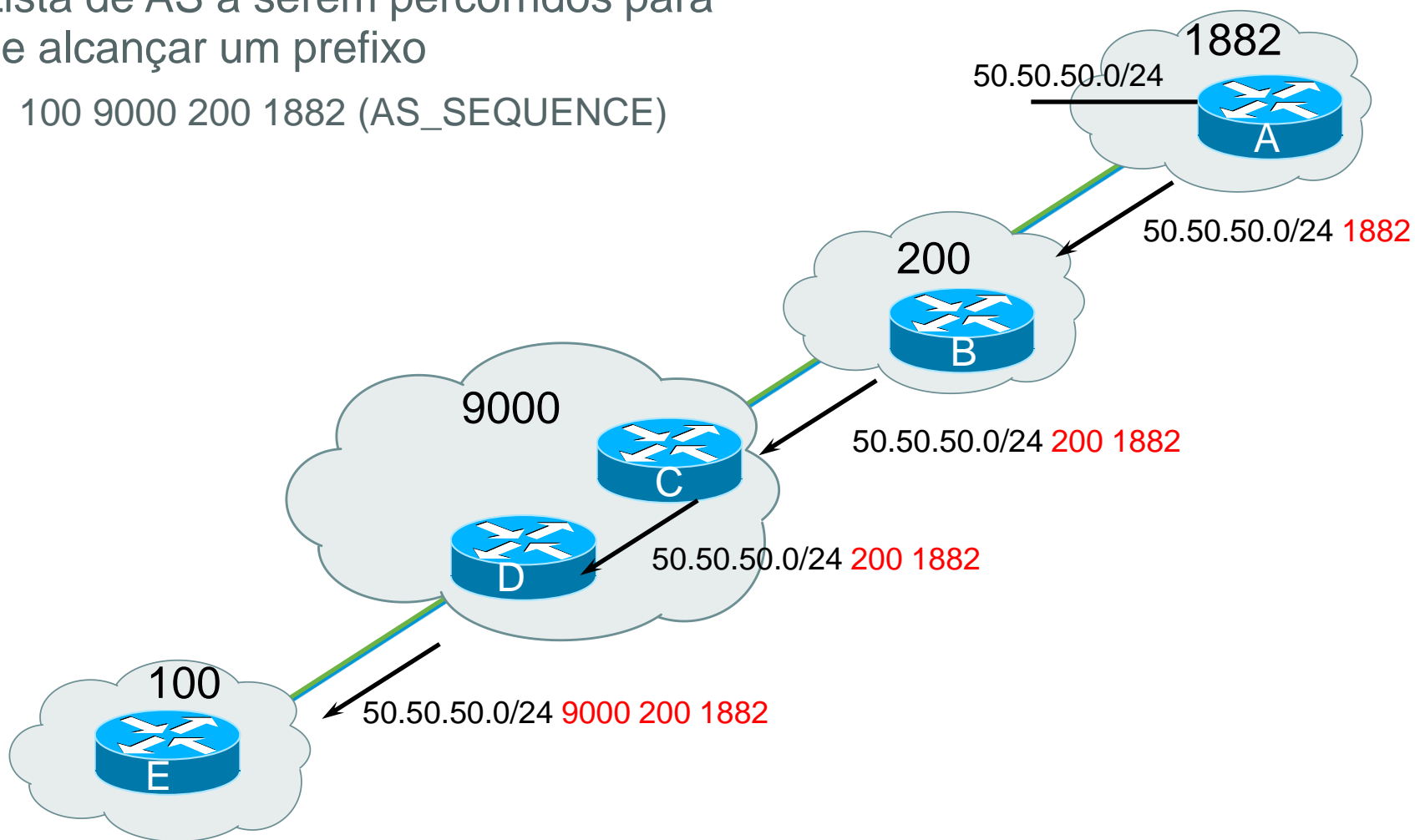


Well-Known
Mandatory

AS_PATH

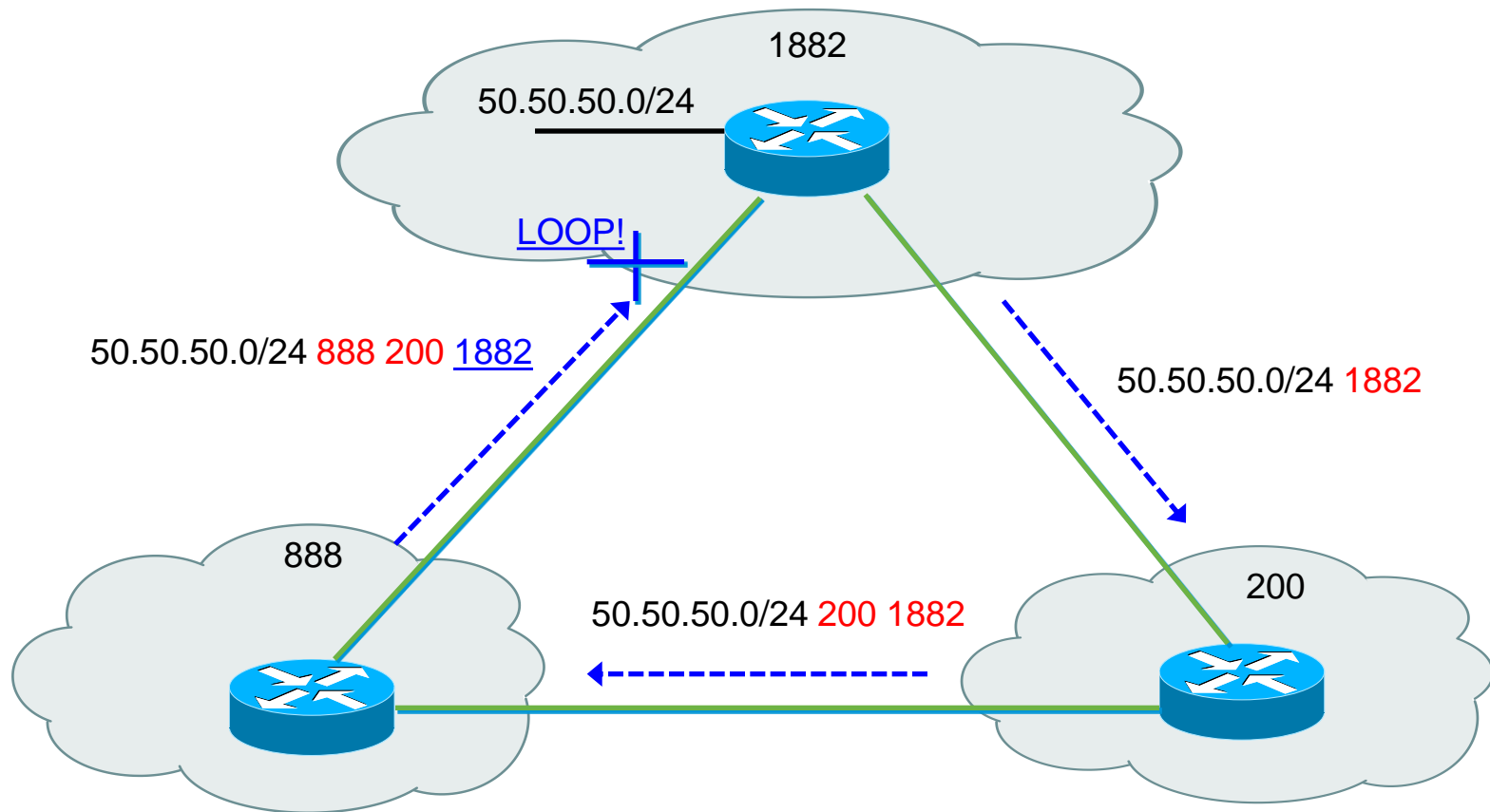
- Lista de AS a serem percorridos para se alcançar um prefixo

100 9000 200 1882 (AS_SEQUENCE)



AS_PATH

- Detecção de Loops



BGP Escolha de Melhor Caminho (Cont.)

7 Menor código ORIGIN:

IGP < EGP < incomplete



Well-Known
Mandatory

ORIGIN

- Origem do prefixo

- Valores:


IGP (i) = via comando network

EGP (e) = aprendido via EGP

incomplete (?) = redistribuição

```
R3#sho ip bgp
BGP table version is 11, local router ID is 2.2.2.3
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
               r RIB-failure, S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete


   Network        Next Hop           Metric LocPrf Weight Path
*> 22.0.0.0        0.0.0.0              0      200    0 7 i
*                  7.7.7.7              0          0 2 i
*                  2.2.2.2              0          0 2 i
*> 33.0.0.0        2.2.2.2              0          0 2 ?
```



BGP Escolha de Melhor Caminho (Cont.)

7 Menor código ORIGIN

8 **Menor Multi-Exit Discriminator (MED)** ... default é 0.
Default: Apenas se os caminhos são recebidos do mesmo AS



Optional
Non-
Transitive

Multi-Exit Discriminator (MED)

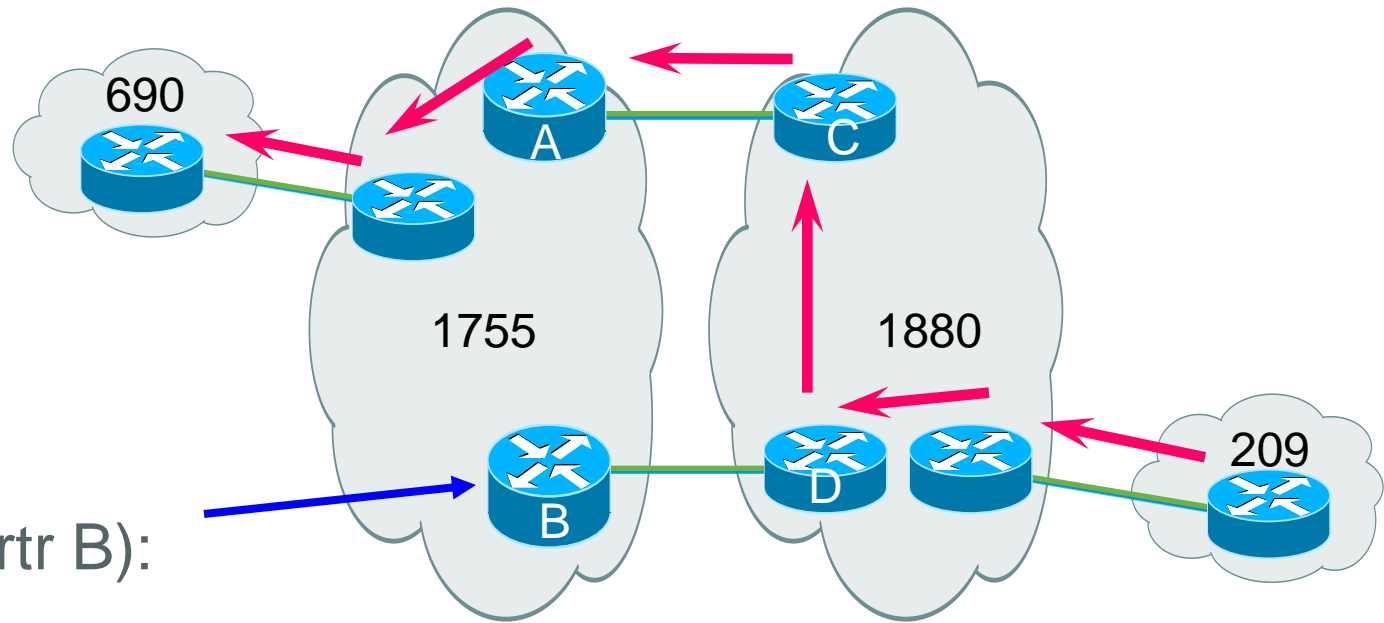
Indicação (para vizinhos externos) do caminho preferido de entrada em um AS

- Utilizado em AS com múltiplos pontos de entradas
- non-transitive

Por default comparado somente entre rotas de um mesmo AS, a não ser que seja configurado `bgp always-compare-med`

MED menor é preferível (default = 0)

MED

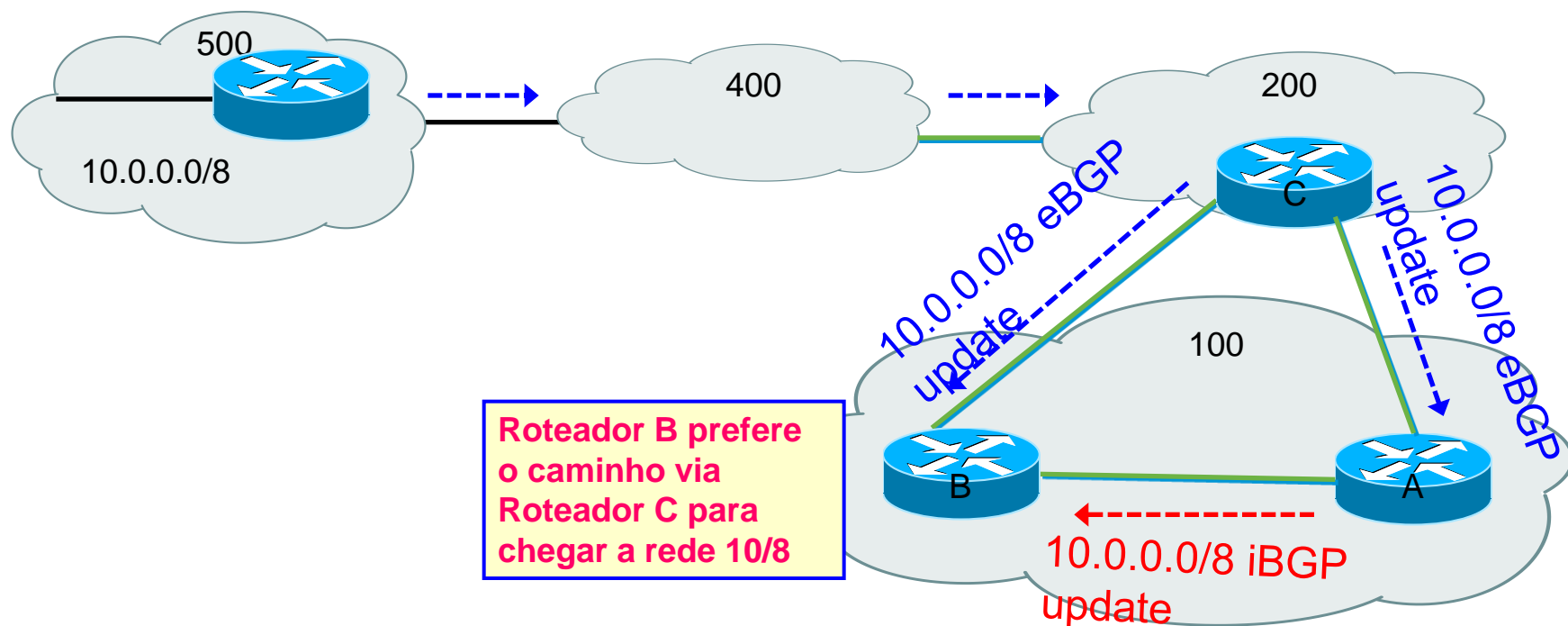


- Configuração (rtr B):

```
router bgp 1755
neighbor x.x.x.x remote-as 1880
neighbor x.x.x.x route-map set_MED out
!
route-map set_MED permit 10
match as-path 2
set metric 2
!
ip as-path access-list 2 permit _690$
```

BGP Escolha de Melhor Caminho (Cont.)

- 7 Menor código ORIGIN:
- 8 Menor Multi-Exit Discriminator (MED)...
- 9 Preferir um caminho *Externo* sobre *Interno*



BGP Escolha de Melhor Caminho (Cont.)

7 Menor código ORIGIN:

IGP < EGP < incomplete

8 Menor Multi-Exit Discriminator (MED)...

9 Preferir um caminho *Externo* sobre *Interno*

10 Menor métrica IGP para o NEXT_HOP

11 Para caminhos eBGP:

- SE multipath está habilitado, o roteador pode instalar até N caminhos paralelos na tabela de roteamento
- Selecionar o mais Antigo a não ser que os updates tenham chegado ao mesmo tempo.
- SE os updates chegaram ao mesmo tempo, ver próximo bullet

BGP Escolha de Melhor Caminho (Cont.)

12 Menor Router-ID

- Originator-ID é considerado para rotas refletidas

13 Menor Cluster-List

- Cliente **deve** estar ciente de atributos dos RR!

14 Menor endereço IP do vizinho

Terceira Pergunta

Qual dos atributos abaixo são mandatórios em um update BGP?

- a) NEXT-HOP
- b) ORIGIN
- c) AS_PATH

Referências

- BGP Best Path Selection Algorithm
http://www.cisco.com/en/US/partner/tech/tk365/technologies_tech_note09186a0080094431.shtml
- BGP Case Studies
http://www.cisco.com/en/US/tech/tk365/technologies_tech_note09186a00800c95bb.shtml
- Cisco Networkers 2004 – Deployment and Analysis of BGP
<http://www.cisco.com/networkers/nw04/presos/sec.html>
- Cisco Support Community – Português
<https://supportforums.cisco.com/community/portuguese>

Perguntas e Respostas



Queremos sua opinião!

Para fazer a avaliação, por favor, clique no endereço fornecido no chat ou no pop-up quando o evento terminar.

Façam suas perguntas!

Utilizem o painel de perguntas e respostas (Q&A) para perguntar aos especialistas agora. Eles já começaram a responder.



Evento Pergunte ao Especialista com Marcio Ferreira

Se você quiser tirar mais dúvidas com o nosso especialista, ele estará respondendo a perguntas entre os dias 18 e 28 de junho neste link:

<https://supportforums.cisco.com/thread/2223362>

O vídeo e as perguntas e respostas serão disponibilizados ainda esta semana no link:

<https://supportforums.cisco.com/community/portuguese/canto-dos-especialistas/webcasts>



Convidamos você a participar da CSC em português e em nossas redes sociais

<https://supportforums.cisco.com/community/portuguese>



Portugal: <http://www.facebook.com/ciscoportugal>

Brasil: <http://www.facebook.com/CiscoDoBrasil>



Portugal: <https://twitter.com/CiscoPortugal>

Brasil: <http://twitter.com/CiscoDoBrasil>



Portugal: <http://www.youtube.com/user/ciscoportugal>

Brasil: <http://www.youtube.com/user/ciscoDoBrasilTV>



Portugal: <http://ciscoportugalblog.wordpress.com/>

Muito Obrigado
por assistir.

Por favor complete o formulário de avaliação e dê
sugestões de temas para os próximos webcasts!



Thank you.

